CORE
Metadata, citation and similar papers at core.ac.uk
Provided by Elsevier - Publisher Connector

Enantioselective Substrate Binding in a Monooxygenase Protein Model by Molecular Dynamics and Docking

K. Anton Feenstra,* Karin Hofstetter,‡ Rolien Bosch,* Andreas Schmid,†‡ Jan N. M. Commandeur,* and Nico P. E. Vermeulen*

*Division of Molecular Toxicology, Department of Pharmacochemistry, Leiden/Amsterdam Center for Drug Research (LACDR), Vrije Universiteit, Amsterdam; †ISAS-Institute for Analytical Sciences, Dortmund, Germany; and ‡Department of Biochemical and Chemical Engineering, University of Dortmund, Dortmund, Germany

ABSTRACT The two-component flavoenzyme styrene monooxygenase (SMO) is an efficient alternative to several chemical epoxidation catalysts on a preparative scale. A first homology model of the catalytic domain (StyA) of SMO was constructed (Protein Data Bank ID 2HD8) based on the structure of *para*-hydroxybenzoate hydroxylase. The StyA protein structure was optimized by restrained molecular dynamics to reproduce specific pre-S binding orientations of styrene. Effects of all 10 point mutations examined were explained by the distance of the site to the styrene and FAD binding sites. Thirteen of 20 ligands could be accommodated in a catalytically active binding orientation, and predicted affinities correlated well with experimental turnover and inhibition. The binding cavity is almost completely hydrophobic except for a hydrogen-bonded network formed by three water molecules, the backbone of residues 300–302, and the flavin ribityl, similar to P293, and three crystal waters in *para*-hydroxybenzoate hydroxylase suggest that P302, T47, and the waters in StyA are a vital component of the catalytic mechanism. The current optimized and validated StyA model provides a good starting point for elucidation of the structural basis of StyA ligand binding and catalysis. Novel insights in the binding of ligands to SMO/StyA, provided by the current protein model, will aid the rational design of mutants with specific, altered enantioselective properties.

INTRODUCTION

Recently, interest in the class of two-component flavoenzyme hydroxylases has emerged (1,2). Two-component flavoenzyme hydroxylases consist of a smaller reductase component that uses NAD(P)H to reduce flavin and a larger hydroxylase component that uses the reduced flavin to perform the hydroxylation reaction. Styrene monooxygenase (SMO) from *Pseudomonas* sp. VLB120 is such a two-component flavoenzyme, consisting of a larger oxygenase domain StyA and a smaller reductase domain StyB, each with an independent catalytic center (2–4). StyA is not dependent on its native reductase; also, different reductases, chemical mediators that generate reduced flavin, and even direct electrochemical reduction on an electrode can be used as electron donor (3,5). The SMO enzyme is highly enantioselective, producing >99% enantiopure *S*-styrene oxide. Of further interest is the development of methods to rationalize and design mutants with altered substrate and/or enantioselectivity (6), which can provide a basis to explore possibly interesting applications in the chiral synthesis of epoxide building blocks for pharmaceuticals or agrochemicals (1). For this purpose, a structural model of the protein based on crystallographic data or from homology model building is indispensable. The class of flavoenzymes is well studied (2,7,8). For 50 FAD-containing oxidoreductase/hydroxylases there are x-ray structures in the Protein Data Bank (PDB) (9), but for SMO no x-ray structures have yet been published.

The current article describes a newly constructed, optimized, and validated homology structure of the StyA oxygenase component of SMO, deposited in the PDB under ID 2HD8 (9). Some general features of the binding cavity were described before, based on a preliminary StyA model (5), but that model was not capable of accommodating substrates in a catalytically active position, nor could it be used to describe or explain substrate binding in detail. For further optimization and for validation of a StyA protein model, experimental data on binding affinity and enzyme turnover for a range of substrates and for selected site-directed mutants were used. The construction and analysis of the current StyA structure comprise a vital step in the rationalization of activity and (enantio)selectivity of catalysis by StyA wild-type (WT) and mutants.

METHODS

Sequence alignment

The primary sequence of StyA (accession code O50214) was obtained from the Sequence Retrieval System accessible from the EMBL (www.embl-heidelberg.de). The search for homologous sequences with resolved crystal structure was carried out using BLAST (10), yielding *para*-hydroxybenzoate hydroxylase (pHBH), PDB identifier (9) 1PBE (11), and 2-phenol-hydroxylase (PhHy), PDB identifier 1FOH (12). A multiple sequence alignment (MSA) of StyA with pHBH and PhHy was manually made based on conserved sequence motifs and substrate recognition sites (SRSs) in flavoproteins (see Table 1) (13), with the homology module of InsightII

Submitted May 9, 2006, and accepted for publication July 12, 2006.

Address reprint requests to N. P. E. Vermeulen, Division of Molecular Toxicology, Dept. of Pharmacochemistry, Leiden/Amsterdam Center for Drug Research (LACDR), Vrije Universiteit, De Boelelaan 1083, 1081HV Amsterdam, The Netherlands. Tel.: 31-20-5987590; Fax: -31-20-5987610; E-mail: NPE.Vermeulen@few.vu.nl.

© 2006 by the Biophysical Society

0006-3495/06/11/3206/11 $2.00

doi: 10.1529/biophysj.106.088633

TABLE 1 Conserved sequence motifs in flavoproteins used for aligning StyA to pHBH

Motif	StyA	Location	Description	Reference
GxGxxG	9–14	βA3–αH1	Postulated structural function for interaction adenosyl-ribose and β-α-bend of terminal $\beta\alpha\beta$	(29,36)
DG	170–171	loop βA4-αH7	Only : ''G'' conserved in StyA	(13)
GDx_6P	294–302	βD3	Putative dual function in FAD/NAD(P)H binding in flavoprotein hydroxylases	(13)
$GxNx_8L$	307–318	αH10	Conserved sequence motif	(13)

(Biosym, San Diego, CA). Further alignment was done on consensus using LOOPP (14). The MSA was refined using data on amino acids relevant for the binding of FAD (Table 2) (11,15–17). Ten different 3D strucures of StyA were generated with the restraint-based comparative modeling program Modeller 4.0 (18). From the structures generated by the Modeller program, the one with the best loop conformations and stereochemical parameters as determined by ProCheck (19) was selected for further modeling.

Protein structure optimization

An optimization scheme employing steepest descents energy minimization (EM) and molecular dynamics (MD) simulations with positional restraints applied to different sets of atoms was used, which could be described as a ''controlled release'' method. The Gromacs MD package, version 3.1.4 (20,21), and the GROMOS 43a1 forcefield (22,23) were used with a 0.8/1.2-nm cutoff and a neighbor list update frequency of five steps. The LINCS algorithm (24) was used to constrain the length of all covalent bonds, and a time step of 2 fs was used. The protein structure, including the substrate styrene and cofactor FAD, was energy minimized in vacuum, first with harmonic position restraints on all nonhydrogen atoms with a force constant of 1000 kJ/mol/nm and then without position restraints. Subsequently, preequilibrated simple point-charge (SPC) water was added in a periodic dodecahedral box with a minimum distance of 1.2 nm between the protein and box edges, and the system was minimized again. Starting velocities were randomly generated from a 300 K Maxwell distribution. Temperature was maintained at 300 K and pressure at 1 bar by weak coupling to a bath using relaxation times of 1 ps and 0.1 ps, respectively.

During a 1-ps MD run the water was allowed to relax while the whole protein was position restrained with a force constant of 1000 kJ/mol/nm. In a series of short MD runs, subsequent parts of the protein were allowed to relax by removing the position restraints for the appropriate atoms. First, the side chains only of the residues not involved in either FAD or styrene binding were released for 1 ps. Next, also the backbone of the binding residues was released for 10 ps. During this whole procedure the positions of both FAD and styrene were also restrained. Finally, only the backbone atoms of the binding residues were restrained for 100 ps, while the rest of the protein as well as FAD and styrene were ''free''. The set of binding residues was defined as all residues that had at least one atom closer than 6 Å to one atom of either FAD or styrene in the nonoptimized structure. These were Ala^{13}, Glu^{38}, Tyr^{39}, Arg^{43}, Thr^{47}, Val^{48}, Glu^{113}, Ala^{209}, Val^{211}, Leu^{220}, Val^{222}, Ala^{224}, Arg^{233}, Phe^{235}, Asp^{295}, Pro^{302}, Gly^{305}, Ala^{308}, Asn^{309}, and Phe^{409} and include all known FAD-interacting residues, as listed in Table 2, with the exception of Asp^{33}. Several 1-ns-long MD simulations without position restraints were performed with styrene and FAD bound in the StyA protein: six started from the nonoptimized structure and six from the optimized structure, each with a different starting velocity randomly generated from a 300 K Maxwell distribution. Dynamic stability of the structure before and after the procedure was assessed by determining the atomic position root mean-square deviation (RMSD) with respect to the starting structure, averaged over the final 0.5 ns of each simulation, and differences in dynamic behavior are assessed from ED projections of the trajectories on the first two eigenvectors of all trajectories combined (25,26).

Candidate model structures were extracted from the 1-ns MD simulations started from the optimized structure by taking average structures over the first 100 ps and over the last 100 ps of the trajectories and energy minimizing to remove unphysical geometries that might result from the averaging, and the substrate styrene was removed. This resulted in a total of 12 candidate structures. Subsequently the substrate styrene was docked into each of the candidate structures. The optimal candidate was selected based on the possibility of reactivity of the docked conformations as judged from the distance to the reactive carbon atom $^{F}C_{4A}$ in the FAD cofactor, the correct prochirality of the docked orientations of styrene as judged from the angle between the styrene plane and that of the isoalloxazine ring of FAD, and the specificity of the docked conformations as judged from the RMSD between different docked conformations within one structure.

StyA mutant structures

Structural models of 14 single (point) mutants of StyA were generated: V274Y/C/H, P275H/A, P302A/H, A298H, F409S, T12P/H, Q341R, L45C/H, and F173H. From the WT StyA structure, the selected amino-acid side chains were replaced by the mutant forms using the maximum overlap principle (retaining all overlapping atoms between WT and mutant residue) (27). The resulting structures were energy minimized to remove unfavorable interactions between the modified side chains and the surrounding protein.

TABLE 2 FAD interacting residues in pHBH with correspondingly aligned residues in StyA, and secondary structure elements

pHBH	StyA	Location	Postulated function	Reference
Ser^{13}	Ala^{13}	βA3-αH1	NH of Ala^{13} H-bonded with $^{F}O_2$ (Pi of FAD), cf. Ser^{13} in pHBH	(11)
Glu^{32}	Asp^{33}	βA2	Conserved D/E. Stabilizes FAD	(16)
Arg^{44}	Arg^{43}	αH2-βE1	H-bonded with $^{A}O_{P1}$ of FAD	(13)
Gln^{102}	Glu^{113}	αH6	H-bonded with $^{F}O_B$ of FAD	(11)
Arg^{220}	Arg^{233}	βB5	Stabilizes out position of FAD in pHBH	(17)
			Binding of FAD_{red} in StyA (putative)	(38)
Asp^{286}	Asp^{295}	βA5-βD3	H-bonded with O-F_1 ($^{F}O_G$), $^{F}O_{P2}$ of FAD	(11)
Pro^{293}	Pro^{302}	βD3-αH10	Conserved, structural; postulated catalytic role in stabilizing transition state of hydroxylation of PhHy (P364)	(15)
			P302 mutants have reduced stability	(5)
Leu^{299}- Asn^{300}	Ala^{308}-Asn^{309}	αH10	Amides H-bonded to flavin ring	(11)

FAD atoms are indicated according to Gatti et al. (17).

Substrate binding conformation prediction

Twenty substances (styrene, 1,2-dihydronaphthaline, 3/4-bromo/chloro/methyl/nitro-stryrene, α/β-methyl-styrene, indene, 1-phenyl-1-cyclohexen, 2/4-vinylpyridine, allylbenzol, 1,3-cyclo-hepta-/octadiene, naphthalene, vinylcyclopentane) were docked into the StyA protein structure using the automated docking program Gold (28). Of each compound, 50 docked conformations were generated and analyzed in terms of ''active'' binding, i.e., at the right distance and orientation with respect to the reactive carbon atom $^{F}C_{4A}$ of FAD. A maximum distance of 6.5 Å was taken, based on $^{F}C_{4A}$-O_p-O_d bond lengths and estimates of distances in the reaction coordinate (15). The prediction of active binding was correlated with measured activities of these substrates to validate details of the active site geometry and chemical composition as well as the bound orientation of the FAD cofactor (29).

Substrate binding affinity prediction

The binding free energy (ΔG) of styrene and several of the other substrates for the StyA WT and mutant proteins were calculated using the linear interaction energy (LIE) approximation (30,31) as follows,

$$\Delta G_{\mathrm{Calc}} \cong \alpha(E_{\mathrm{VdW,bound}} - E_{\mathrm{VdW,free}}) + \beta(E_{\mathrm{Coul,bound}} - E_{\mathrm{Coul,free}}), \quad (1)$$

where $E_{\mathrm{VdW,bound}}$ and $E_{\mathrm{Coul,bound}}$ are the van der Waals and the Coulomb interaction energies, respectively, between the substrate and its surroundings in the bound state, i.e., protein, FAD, and water, and $E_{\mathrm{VdW,free}}$ and $E_{\mathrm{Coul,free}}$ between substrate and surroundings in the free state, i.e., only water. These energies are obtained from MD simulations of the respective states, and error margins of the calculated ΔG are obtained from the respective energy fluctuations during the simulations. Using the solvent-accessible surface area of styrene calculated using the program MSMS (32) and the method proposed by Wang et al. (33) based on the correlation between weighted nonpolar desolvation ratio (WNDR) and α, an α of 0.98 was obtained. For the calculation of binding affinities for uncharged ligands without hydroxyl groups, the suggested value of 0.43 for β is used (34). From the free energy of binding (ΔG_{Calc}), the binding affinity is calculated as follows,

$$K_{\mathrm{d,Calc}} = \exp(\Delta G_{\mathrm{Calc}}/k_{\mathrm{B}}T), \quad (2)$$

where k_{B} is Boltzmann's constant and $T = 300$ K.

Selected docked conformations according to the criteria mentioned above of styrene, 3-/4-/α-/β-methylstyrene, 2-/4-vinylpyridine, vinylcyclopentane, 1,2-dihidronaphthaline, indene, and naphthalene, for the StyA-substrate complexes of WT and for styrene in the WT and the F235A/S, V274C, P275A/H, P302A/H, and F409S mutants were energy minimized, solvated, and equilibrated as described under Protein structure optimization. Four hundred picoseconds of production MD simulations were performed, and average interaction energies were collected over the final 300 ps. For interaction energies of the free substrate in solvent, the same procedure was followed, with 100 ps of production MD simulation and energies averaged over the last 50 ps.

RESULTS

StyA sequence alignment

A homology model of StyA (SMO catalytic domain) was constructed based on the x-ray crystallographic structure of pHBH, the closest homolog to StyA with 23% pairwise sequence identity. An overview of this protein model is presented in Fig. 1 *A*. The higher local similarity with PhHy (pairwise identity 22%) was also used for the MSA. Conserved motifs are summarized in Table 1. The conserved sequence motif GxGxxG (StyA residues 9–14) is thought to play an important role in the interaction with the ribose of the adenosyl moiety of FAD in pHBH (35) and in FAD-binding proteins in general (29,36). It is localized in the structurally conserved β-α-β unit at the C-terminal end of hydroxylases and monooxygenases (13). The short DG motif is located in loop βA4-αH7 of the FAD-binding domain of pHBH and is situated near the cleft leading toward the active site (13); in StyA only the Gly^{171} is conserved. The GDx_6P motif (StyA residues 294–302) is conserved in flavoprotein hydroxylases with a putative dual function in FAD and/or NAD(P)H binding (13) and is located in strand βD3 (note that StyA binds FAD, not NAD(P)H). The conserved $GxNx_8L$ motif (13) corresponds to StyA residues 307–318 and is located in helix αH10. Further alignment was done on consensus using the LOOPP program output (14).

The MSA was refined by comparing the FAD binding residues in pHBH (11,13,15,16) (see Table 2) with residues in StyA. Ala^{13}, Asp^{33}, Arg^{43}, Glu^{113}, Arg^{233}, Asp^{295}, Pro^{302}, Ala^{308}, and Asn^{309} in StyA were found to correspond with Ser^{13}, Glu^{32}, Arg^{44}, Gln^{102}, Arg^{220}, Asp^{286}, Pro^{293}, Leu^{299}, and Asn^{300} in pHBH, respectively. The complete MSA of pHBH, PhHy, and StyA is presented in Table 3. In this final MSA between StyA and pHBH, 173 (38%) are strongly conserved (conservation score $\geq$8, e.g., Val-Ala, see Table 3).

In Table 4, the effects of the single amino-acid substitutions V274Y/C/H, P275H/A, P302A/H, A298H, F409S, T12P/H, Q341R, L45C/H, F173H in the StyA protein on the activity of the enzyme are listed as well as the closest separation between any of the atoms of the corresponding amino acids in the model structure and styrene and FAD. These respective residues are shown in Fig. 1 *B* in relation to the location of styrene and FAD in the model. For all residues closer than 5 Å to either styrene or FAD (T12, L45, V274, P302 and F409), a large effect of at least one of the substitutions (<50% activity versus wild-type) was observed. Conversely, mutations in residues at more than 5.5 Å from styrene and FAD (F173, P275, A298, and Q341) showed small effects of all substitutions (at least 50% activity versus wild-type). In addition, for mutations in six positions, the stabilization of bound reduced FAD was determined (see Table 4), and it was changed by at least one mutation in five of them (T12, L45, F173, F235, and P302), all closer than 5.5 Å to FAD in the model. With the effects of all 10 mutated residue positions in StyA explained by distances in the StyA model structure, the alignment of the StyA sequence to the template structures pHBH and PhHy is properly validated, especially in the most important areas of the enzyme: the binding sites of styrene and FAD.

StyA structure and optimization

The initial protein model structure was found to be moderately stable in EM and MD simulations (20–23), with Cα atoms RMSD with respect to the starting conformation rising during

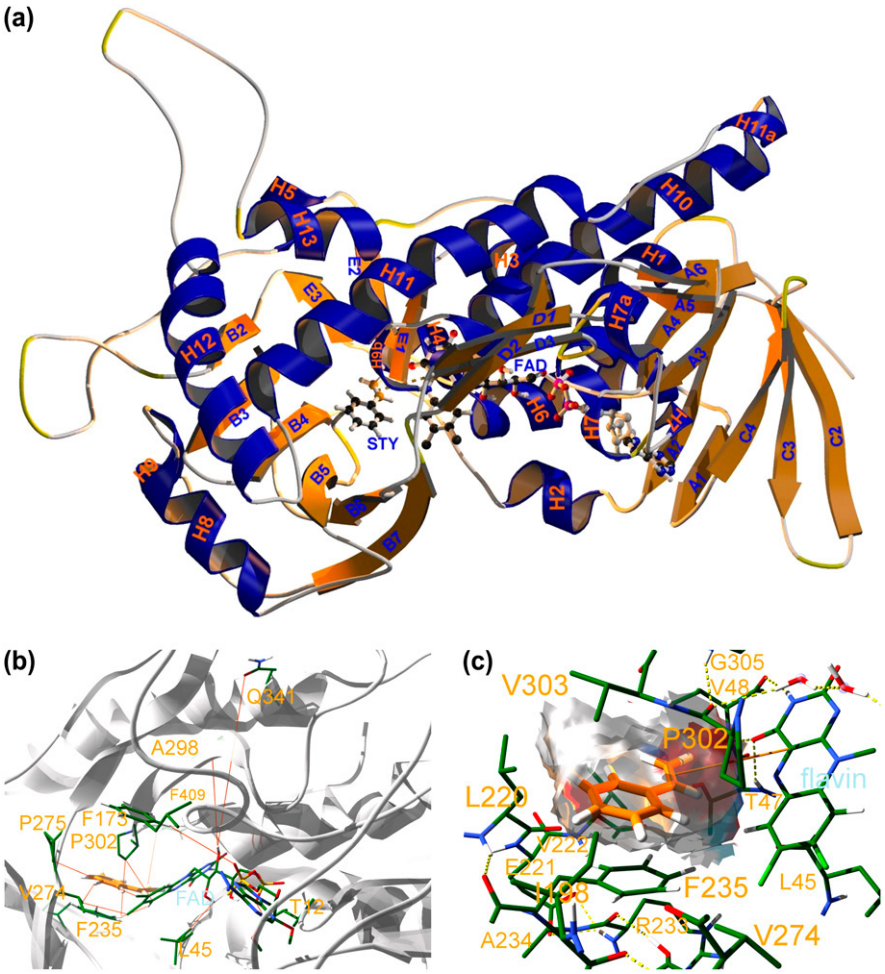
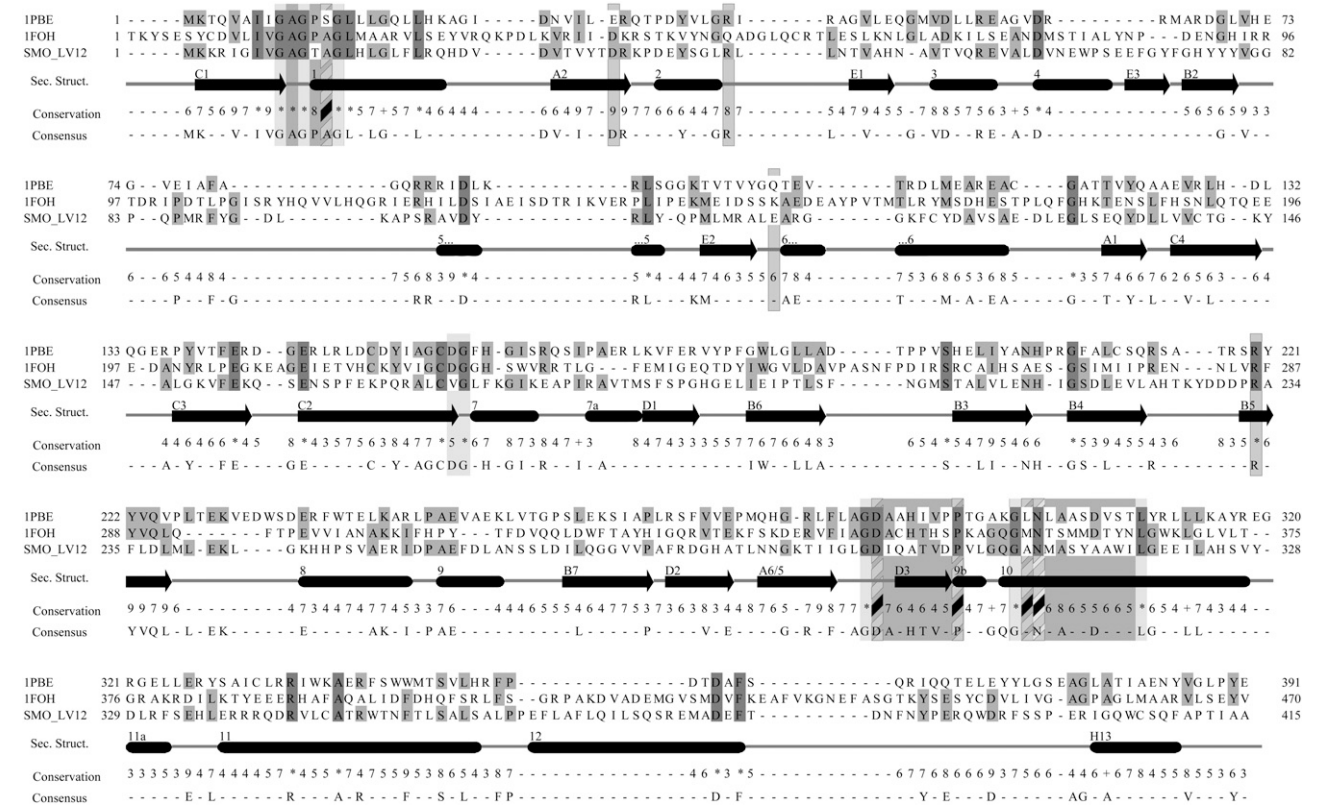


FIGURE 1 StyA, the substrate styrene (STY) and the cofactor FAD. (*A*) Overview of the whole StyA structure. Secondary structure elements are labeled according to the template pHBH (11). (*B*) Mutated residues around the StyA active site described in this study. Orange lines indicate the closest distances between the mutated residues and styrene or FAD. (*C*) Styrene binding in the StyA active site. The styrene binding cavity surface is shown, and the frontmost part is clipped to allow an inside view. Bound waters are highlighted with transparent blue spheres. The catalytic $^{F}C_{4A}$ in the isoalloxazine ring of FAD and the vinyl group of Sty are highlighted in orange; the line in between shows their relative orientation. Hydrogen bonds among FAD, binding residues, and water are indicated with dashed yellow lines.

six 1-ns free MD simulations to plateau values of 4.2 ± 0.5 Å. The Cα atom RMSD between the protein structure before and after the controlled-release optimization was 2.8 Å for the whole protein (3.4 Å for all atoms) and 0.9 Å for the Cαs of the active site only (1.7 Å for all atoms). The optimized protein structure was more stable with corresponding RMSD plateau values of 3.6 ± 0.3 Å during six 1-ns free MD simulations, which indicates a good dynamic stability for a protein of this size and type on a timescale of nanoseconds. The Cα RMSD plateau of the binding residues was reduced from 2.0 ± 0.3 Å in the simulations before optimization to 1.5 ± 0.3 Å in the simulations after optimization. This indicates a good dynamic stability of the active site region. Six candidate model structures were extracted from six simulations of the optimized structure by averaging the conformations of the first 100 ps, and another six structures from the final 100 ps. ProCheck analysis (19) of the distribution of backbone dihedral angles of these candidate structures indicated the structures after the optimization simulations still possessed good stereochemical quality (19): in all cases 94 ± 1% of residues were in core or allowed regions. Table 5 shows a summary of Ramachandran distributions at different stages in the optimization. For the initial model, 78% of main-chain bond lengths were within limits, and two were off the graph; for all other models 100% were within limits. As a general trend, it was observed that, during the five stages of the controlled-release optimization, the quality decreased somewhat; i.e., the Ramachandran plot shifted from 74% core and 20% allowed to 63% and 32%, respectively. However, much of this decrease was made up again by taking the minimized average structure over 100 ps of simulation in the final stage.

In Fig. 2, the essential dynamics (ED) projection (25,26) is shown for the changes in the protein and active site structure on optimization and during the MD simulations. Distances between two points in this plot are closely related to the familiar RMSD measure, but in addition, differences between the trajectories visualize motion in different directions. Motions along the two axes in the plot correspond to the two largest collective modes of motion of the enzyme, also known as ''essential modes'' (25). This illustrates the difference in behavior between the simulations started from the nonoptimized structure and those started from the optimized structure. The stretched appearance of the traces of the nonoptimized simulations is caused by a ballistic type of motion indicative of high internal strain in the protein structure and corresponds to the high RMSD plateau level of 4.2 ±

TABLE 3 Multiple sequence alignment (MSA) among StyA (SMO_LV12), pHBH (1PBE), and PhHy (1FOH) as produced by CLUSTAL W (1.82)



On light and medium gray background the conserved sequence motifs (Table 1) are indicated, and on boxed gray the FAD binding residues based on the crystal structure of pHBH (1PBE) (Table 2). Identical residues are indicated by a dark gray background, lighter if only two of three match. Secondary structure elements (for α-helix and for β-sheet) in the StyA structure are indicated, and labeled according to the pHBH template (11). Conservation scores for the alignment are indicated on a scale from 1 to 9 and + for nearly and * for completely identical, respectively, as calculated by JalView 2.06 (41). Finally, the consensus sequence is shown.

0.5 Å for the whole protein and 2.0 ± 0.3 Å for the active site. In the projection of the whole protein, the difference in the nonoptimized and optimized structures of 2.9 Å can be seen as a shift of the starting point of the corresponding simulations, whereas in contrast, in projection of the active site, the smaller structural shift of 0.9 Å is hardly visible. The optimized structure, in contrast, shows a much more diffusive nature of the projections of the simulations, which is indicative of a protein structure with native-like properties (37) and corresponds to the lower RMSDs of 3.6 ± 0.3 Å and 1.5 ± 0.3 Å for protein and active site, respectively.

Of all optimized StyA protein structures, the predictions from automated docking showed one structure to bind styrene in a single tight cluster with a maximum RMSD between members of the cluster of 0.5 Å, in a catalytically active position with respect to the FAD cofactor reactive center (<6.5 Å between the ethylene of styrene and $^{F}C_{4A}$, see Fig. 1 *C*), and an orientation prochiral to the formation of *S*-styrene oxide exclusively. In the other structures, styrene was bound in a noncatalytically relevant position, in widely varying positions, or in a mixed prochiral orientation. No protein structure was found that would support exclusive pro(*R*) orientation. The one structure showing the correct, tight, and pro(*S*) binding characteristics for styrene was chosen for subsequent validation. The final StyA structure is shown in Fig. 1 *A*, and details of the styrene binding pocket and the bound orientation of styrene are shown in Fig. 1 *C*.

StyA structure validation

Binding free energies (ΔG) for styrene in the wild-type (WT) StyA and the mutants F235A/S, V274C, P275A/H, P302A/H, and F409S were calculated from MD simulations using the LIE method, using Eq. 1 (30,31). The docked orientation of styrene in the StyA WT structure was also used as the starting orientation for the MD simulations in the mutant structures (see Fig. 1, *A* and *C*). Calculated free energies (ΔG) of binding are listed in Table 4. The average distances between the catalytically active carbon $^{F}C_{4A}$ of FAD and the vinyl group of styrene (see Fig. 1 *C*) are listed in Table 4. A distance of <6.5 Å can be considered reactive (15).

In Table 6, experimental activities classified as high (>20%), low (1–20%), or no activity (<1%) relative to styrene for the 20 selected styrene-homologous substrates are

TABLE 4 Mutants of StyA, shortest distances between atoms of the mutated residues and the styrene (Sty) and FAD binding pockets in the protein structure, measured relative activities, and binding affinities of reduced FAD, average distances between $^{F}C_{4A}$ of FAD and the vinyl carbons (CH) of styrene during LIE-MD simulations, and calculated free energies (ΔG_{Calc}) for styrene in selected mutants

	Distances (Å)					
Mutation	Sty	FAD	Rel. act. (%)†	FAD_{red} binding†‡	Aver. dist. $^{F}C_{4A}$-CH (Å)†§	ΔG_{Calc} (kcal/mol)†¶
WT	—	—	100		4.1 ± 0.4	−9.1 ± 1.2
T12H	13	1.8	<1*	≪	—	—
T12P	13	1.8	20	—	—	—
L45C	4.6	2.9	17	>	—	—
L45H	4.6	2.9	11	~	—	—
F173H	13	5.3	<1*	≪	—	—
F235A	2.8	2.6	—	—	6.2 ± 0.7	−9.0 ± 1.9
F235S	2.8	2.6	—	—	7.4 ± 1.3	−8.4 ± 1.5
F235V	2.8	2.6	1	~	—	—
F235Y	2.8	2.6	31	<	—	—
V274C	6.1	3.9	90	—	7.5 ± 0.4	−10.1 ± 1.1
V274H	6.1	3.9	10*	—	—	—
V274Y	6.1	3.9	100	—	—	—
P275A	7.7	6.3	65	—	7.4 ± 0.5	−9.3 ± 1.2
P275H	7.7	6.3	110	—	7.9 ± 0.3	−9.4 ± 1.3
A298H	13	5.9	55	—	—	—
P302A	3.3	4.1	10*	≪	6.5 ± 0.5	−8.8 ± 1.3
P302H	3.3	4.1	<10*	—	9.7 ± 0.3	−12.5 ± 1.3
Q341R	18	13	100	—	—	—
F409S	5.3	5.5	70	~	7.9 ± 0.5	−9.0 ± 1.3

*Inclusion bodies formed.
†''—', 'not determined'.
‡More '>', similar '~', less '<' or much less '≪' than wild type.
§Average distances between FAD catalytic C_{4A} and styrene ethylene atoms during LIE-MD simulations.
¶Error margins derived from energy fluctuations during the MD simulations.

listed. Sixteen of these could be docked at a distance to the catalytically active carbon $^{F}C_{4A}$ and in an orientation with respect to the isoalloxazine ring of FAD, corresponding to an active conformation. However, 3-bromo-, 4-chloro-, and 3- and 4-nitrostyrene were predicted by the docking algorithm to bind in an orientation relative to FAD in which the aromatic rings were stacked.

For styrene, 3-, 4-, α- and β-methylstyrene, 1,2-dihydronaphthaline, naphthalene, indene, 2- and 4- vinylpyridine, vinylcyclopentane, allylbenzol, and 1,3-cycloheptadiene binding free energies (ΔG) were determined using the LIE method and are listed in Table 6. In Fig. 3, it can be seen that in general the substrates with low activity have low predicted binding affinities (≤ -8.7 kcal/mol), whereas most with

TABLE 5 ProCheck Ramachandran distributions and other quality factors for the StyA structures in various stages of optimization and for the final model

	Ramachandran*					MCB angle*†			Planar groups*			G-factors		
Stage	Core	Allowed	Generous	Disallowed	Bad contacts	Within limit	Highlight	Off graph	Within limit	Highlight	Off graph	Dihedrals	Covalent	Overall
Initial model	74	20	4.5	1.7	30	60	40	13	98	2	0	0.46	0.8	0.57
Minimized	77	16	4.7	1.9	2	61	39	3	40	60	28	−0.71	−0.26	−0.49
Solvation	72	23	3.3	1.7	6	49	51	5	24	76	61	−0.97	−1.27	−0.99
Side chain relax	66	28	4.2	1.9	4	52	48	1	30	71	49	−1.03	−1.13	−0.98
Binding site sc	65	30	2.5	1.9	7	51	49	5	28	72	48	−0.92	−1.21	−0.94
Binding site relax	63	32	2.8	2.8	3	52	48	3	31	69	48	−0.86	−1.04	−0.84
Candidates '1st 100 ps'	69 ± 0.7	26 ± 1.3	1.9 ± 0.7	3.1 ± 0.2	1 ± 0.6	76 ± 1.2	24 ± 1.2	0	57 ± 3	43 ± 3	19 ± 1.3	–0.60 ± 0.01	0.075 ± 0.005	–0.31 ± 0.008
Candidates 'Last 100 ps'	67 ± 2.2	27 ± 2.4	2.6 ± 0.7	3.0 ± 0.4	0.8 ± 0.8	74 ± 1.5	26 ± 1.5	0	57 ± 4	43 ± 4	18 ± 4	–0.60 ± 0.02	0.05 ± 0.02	–0.32 ± 0.02
Final model	70	25	1.4	3.3	1	78	22	0	56	44	20	−0.62	0.08	−0.32

For the candidate final models, six from the first and six from the last 100 ps of the simulations (see text for details), average values and standard deviations are listed.
*Percentage within region indicated.
†Main chain bond—angles.

higher activity (''+'' or ''o'') have higher affinity (≥ -8.7 kcal/mol). For the individual substrates, however, the error margins are such that the differences are not significant. Finally, the relatively high inhibition of styrene catalysis by 4- and β-methylstyrene and the relatively low inhibition by α-methylstyrene determined experimentally, are reproduced by the calculated (relative) binding affinities of these substrates (see Table 6). The final StyA structure has been deposited in the PDB under ID 2HD8.

Structural features of styrene, FAD, and water binding in StyA

Styrene binds in the StyA binding pocket stacked with Phe^{235} at the bottom and with Val^{303} at the top, as is shown in Fig. 1 *C*. Leu^{220} and the di-methyl-phenyl ring of FAD make up the left and right sides of the binding pocket, respectively. Finally, Ile^{198} makes up the front, and Val^{222} and the hydroxyl group of Thr^{47} the rear side. Of these six residues, five are at ~3.5 Å from styrene, whereas Val^{303} and FAD are at ~4 Å. Thr^{47} is the only polar side chain in the styrene binding pocket. It is in a position close to the catalytic $^{F}C_{4A}$ of FAD and the ethylene group of styrene, where it may be involved in stabilization of an intermediate in the reaction. Ten additional residues can be found that have at least one atom within 6.0 Å of styrene, shown as well in Fig. 1 *C*, but none of these has any hydrophilic parts within 6 Å of styrene. Furthermore, the backbone carbonyls of Ile^{198}, Leu^{220}, Pro^{302}, and Val^{303} are located inside the pocket. These active site residues and secondary structure elements are summarized in Table 7. In the strongly hydrophobic StyA styrene binding cavity, no water is present at all. A putative substrate access channel can be identified going through the FAD binding site, which is blocked when FAD is bound.

The interactions of FAD with StyA are shown in Fig. 4. The isoalloxazine ring is H-bonded to Thr^{47}, Val^{48}, and Asn^{309}, and additional H-bonds are formed with Ala^{13}, Asn^{46}, and Ala^{308} (not shown). Hydrophobic contacts are made with Gln^{306} and Gly^{307}, and additionally with Leu^{45} and Gly^{305} (not shown). The flavin ribityl hydroxyl groups are H-bonded to the pyrophosphate and to Glu^{113}, Asp^{295}, and some water molecules (not all waters are shown). The pyrophosphate is H-bonded to Thr^{12}, via water to Arg^{34}, and in addition to Arg^{43} (not shown). The adenosine ribose is H-bonded to the backbone of Arg^{34} and additionally to Asp^{33} and the backbone of Ala^{10} and Gly^{11}. Finally, the adenine double ring is H-bonded to Leu^{139} and Lys^{174} (not shown) and makes hydrophobic contacts with Arg^{34}. The front (as in Fig. 4) of the FAD binding pocket of StyA is open and accessible to the solvent; ~16 water molecules hydrate the bound FAD, several of which are worth mentioning. A string of three waters runs from the active site along the backbone carbonyls of Pro^{302}, Asp^{301}, Glu^{306}, and Gly^{307} to the ribityl of FAD. Two waters meditate H-bonds between Arg^{34} and Asp^{295} and the ribityl and pyrophosphate, and three more (not shown) extend this H-bonded network. An additional three waters (not shown) mediate H-bonds between the adenine and Leu^{138}, Leu^{139}, and Glu^{178}.

DISCUSSION

StyA structure building and optimization

We have examined the substrate binding of the oxygenase domain (StyA) of the styrene monooxygenase enzyme from *Pseudomonas* sp. VLB120, in its wild-type form and several site-directed mutants, using a newly constructed homology model. For this purpose, ligand binding orientations and affinities were predicted for the wild-type StyA and 17 mutants and for 20 other substrates for the wild-type StyA. The predictions were used to rationalize the observed relative activities of styrene catalysis determined experimentally.

The StyA protein model was based on the crystallized *para*-hydroxybenzoate hydroxylase (pHBH) and 2-phenol hydroxylase (PhHy). The pairwise identity between pHBH and StyA was high enough (23%) to make homology modeling feasible and was further facilitated by the strong sequence similarity between pHBH and StyA of ~38% (conservation

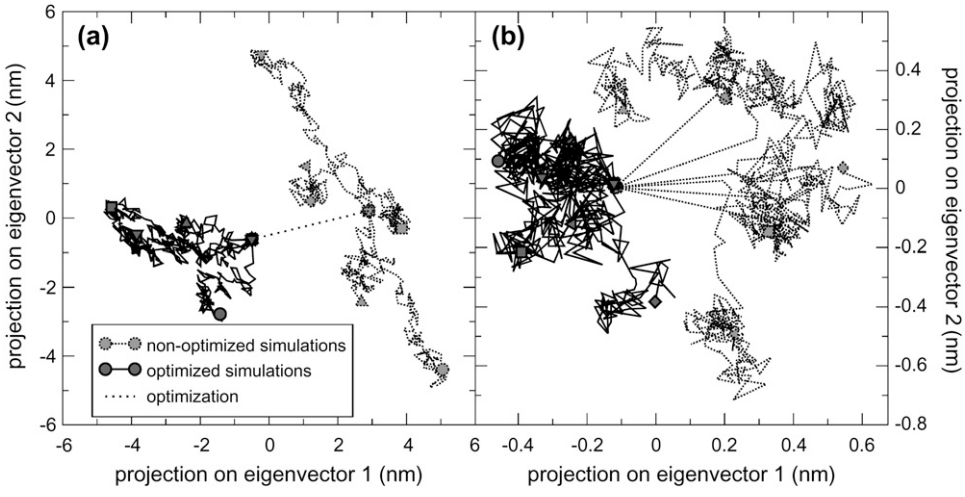


FIGURE 2 MD simulations for optimization and evaluation of optimized structures projected onto the two essential modes (the two largest principal eigenvectors) of (*a*) Cα positions of all 415 residues of the protein; and (*b*) Cα positions of the binding residues only. The dotted lines indicate differences in conformation before and after the controlled-release optimization process. These are prominent for the whole protein but hardly visible for the active site. Note the difference in scale between the plots.

TABLE 6 Substrates and other compounds with known activity in StyA*

Substrate	Docking†			Exp. Act.‡	Calculated affinity§¶		Inhibition§	
	Distance	Orientation	Pred. act.		ΔG (kCal/mol)‖	K_d (nM)	Predicted**	Exp††
Styrene	+	+	+	+	−9.0 ± 1.2	162	—	—
3-Methyl-styrene	–/+	+	+	+	−9.4 ± 2.2	73	—	—
4-Methyl-styrene	+	+	+	+	−10.8 ± 1.4	6.6	25	0.2
α-Methyl-styrene	+	+	+	+	−8.7 ± 1.2	242	0.7	0.08
β-Methyl-styrene	+/+	+/+	+/+	+	−10.8 ± 1.2	6.6	25	1.5
3-Bromostyrene	–/+	–	–	+	–	—	—	—
4-Bromostyrene	+	+	+	+	–	—	—	—
3-Chlorostyrene	+/+	–/+	+	+	–	—	—	—
4-Chlorostyrene	–	–	–	+	–	—	—	—
3-Nitrostyrene	+/+	–/–	–	+	–	—	—	—
4-Nitrostyrene	+	–	–	+	–	—	—	—
1,2-Dihydro-naphthalene	+/+	+/–	+	+	−10.6 ± 1.6	9.8	—	—
Indene	+	+	+	+	−10.1 ± 1.4	22	—	—
1-Phenyl-1-cyclohexen	+	+	+	o	–	—	—	—
2-Vinylpyridine	+/o	+/–	+/o	o	−11.5 ± 2.9	2.0	—	—
Allylbenzol	+	+	+	o	−8.7 ± 1.7	242	—	—
1,3-Cyclo-heptadiene	+	+	+	–	−8.3 ± 1.4	539	—	—
4-Vinylpyridine	+	+	+	–	−9.9 ± 2.9	33	—	—
cis,cis-1,3-Cyclooctadiene	+	+	+	–	–	—	—	—
Naphthalene	+	+	+	–	−8.7 ± 1.4	242	—	—
Vinylcyclopentane	+/+	+/+	+/+	–	−8.1 ± 1.2	805	—	—

*Distances, orientation, and predicted activity (pred. act.) from docking, and measured activities (Exp. Act.), calculated binding affinities (ΔG and K_d cf. Eq. (2)) and predicted (pred) and experimental (exp) inhibition from LIE are listed.

†Binding distance (dist) and orientation (orient) from automated docking: active (+), intermediate (o) and nonactive (−). Alternate docking solutions are separated by a slash (/).

‡Experimental versus predicted styrene activity (pred, i.e., 'active' binding distance and orientation) and experimental (exp): high ('+', >20%), low ('o', <20%) or no ('−', <1%).

§'—', 'not determined'.

¶Binding free energy ΔG (kcal/mol) and affinity K_d (nM) calculated using the LIE method.

‖Error margins derived from energy fluctuations during MD simulations.

**Relative inhibition versus styrene calculated as $K^{Sty}_{d,Calc}/K_{d,Calc}$.

††Inhibition measured experimentally as relative rate of epoxidation versus styrene at equimolar concentrations of 1 mM.

score ≥8 in the MSA of Table 3). Strongly conserved structural and functional features within the oxidoreductase enzyme family can be traced down to several key patterns of conservation (Table 1) (13,29), certainly for the most important active site region of the enzyme (Table 2) (11,15–17,38), lending additional reliability to the alignment. All known pHBH active site residues were very strongly conserved throughout the MSA (conservation score ≥8 in Table 3), and their counterparts in the StyA sequence (Table 2) could be incorporated in the active site region of the StyA protein structure (Fig. 4). For several additional residues that are known to influence the StyA catalytic function (5), their close proximity to the substrate and/or cofactor is shown (Table 4).

This first StyA homology model has been demonstrated to be stable in MD simulations (Fig. 2). For the whole StyA protein, the change in structure from the nonoptimized to the optimized structure is rather large (2.9 Å RMSD on Cα), whereas for the active site only, the concomitant structural shift is relatively small (1.7 Å RMSD on Cα). In addition, from the ED projections (Fig. 2), it can be seen that the structural changes during free MD simulation of the nonoptimized structure are ballistic and indicative of high levels of internal strain in the protein structure, in contrast to the simulations of the optimized structure, which are diffusive, which is a characteristic of stable, native-state protein structures (37). These are clear effects of the controlled-release optimization method employed that minimizes structural changes in the active site region while

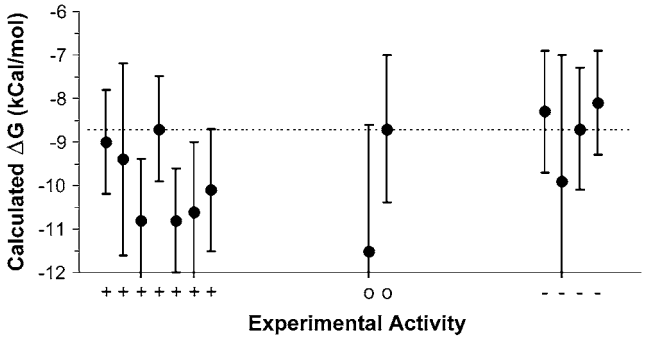


FIGURE 3 Calculated binding free energies (ΔG_{calc}) for 13 of the substrates, as listed in Table 6, following the ordering and classification of experimental activities as in Table 6.

TABLE 7 Active site residues with at least one atom within 6 Å from at least one atom of the bound styrene*

Sec. struct. element	Active site residues	Location in active site†
βE1	Leu[45], **Thr[47]**, Val[48]	Rear
βB6	**Ile[198]**, *Thr[200]*	Front bottom, access channel
βB4	**Leu[220]**, Glu[221], **Val[222]**	Bottom left
βB5	*Arg[233]*, *Ala[234]*, **Phe[235]**	Bottom (center)
βB7	Val[274]	Front right, access channel
αH9b	Pro[302], Val[303], Gly[305]	Top, access channel
Cofactor FAD		Right

*Location in secondary structure elements and in space around the binding pocket for the StyA protein model are indicated. For completeness the FAD cofactor is also listed. Five residues with atoms closer than 3.5 Å constitute the main binding residues and are marked in bold; the three that are not actually at the pocket surface are marked in italics.
†See Fig. 1 *C*.

allowing the rest of the protein structure to relax significantly.

StyA structure validation with experimental data

For the 13 mutants of StyA studied (Table 4 and Fig. 1 *B*), a clear correlation was observed between the distance of a mutation site to the closest of the styrene and FAD binding sites and the maximum effect of the substitutions on the measured enzyme activity, indicating a good overall alignment of StyA with the template structure. The P302 substitutions are particularly worth mentioning. The P302H mutant has a low measured activity for styrene (<10% versus WT), and styrene is predicted to bind very tightly ($\Delta G < -11$ kcal/mol, $K_d < 1$ nM) but too far away for catalysis (~10 Å) (see Table 4). It appears that the histidine at position 302, which is located just above the isoalloxazine ring of FAD, may block access to the catalytic site in StyA (see Fig. 1 *C*). Alternately, the P302H mutation could affect the structure or stability of the βD3-αH10 loop. The P302A mutant has a low measured activity (10%), but the predicted binding affinity of styrene is comparable to WT ($\Delta G \approx -9$ kcal/mol, $K_d \approx 100$ nM), and it binds close enough for catalysis (~6.5 Å). Possibly, changing the stiff proline for the more flexible alanine adversely affects the protein's stability, as is indeed also seen from the increased appearance of inclusion bodies (5). In the V274C mutant, styrene is predicted to bind strongly ($\Delta G = -10.1$ kcal/mol, $K_d = 24$ nM), but at too large distance (>7 Å) to be actively metabolized (Table 4).

All 21 compounds studied (Table 6) could be docked into the active site cavity. For 13 compounds (Table 6), affinities were predicted using the LIE method. As a trend, the active compounds showed a higher predicted affinity, although for individual cases the error margins are such that significance is minimal. In addition, it must be borne in mind that in other cases too, the actual binding affinity could well be the limiting factor for efficient turnover. Only 3-bromo-, 4-chloro-, and 3- and 4-nitrostyrene were predicted to bind in an orientation unsuited for metabolism, whereas they were nevertheless metabolized in our experiment. This might be attributed to the chemical nature of the relatively large, electron-rich, and polarizable bromo-, chloro- and nitro- substituents, whose properties are difficult to describe appropriately in a force field or scoring function. The two substrates with lowest predicted affinity (>250 nM; cycloheptadiene and vinylcyclopentane) lack the characteristic aromatic six-membered ring and are also experimentally inactive. The other substrate that lacks the aromatic ring is cyclooctadiene, which, although predicted to bind in the right place and at good affinity, is not catalyzed (Table 6). This leads to the tentative conclusion that the aromatic six-membered ring might be important for binding and essential for catalysis. The final StyA structure has been deposited in the PDB under ID 2HD8.

The StyA active site and styrene and FAD binding regions

A close look at the styrene binding cavity and active site revealed that the hydroxyl group of Tyr[47] and carbonyl group of Pro[302] are at ~4 Å from the catalytic $^{F}C_{4A}$ of FAD and the ethylene group of styrene in an otherwise completely hydrophobic binding cavity (at the rear and front, respectively, in Fig. 1 *C*). The corresponding Pro[293] in pHBH (Table 2) is implicated in stabilization of intermediate state(s) of the

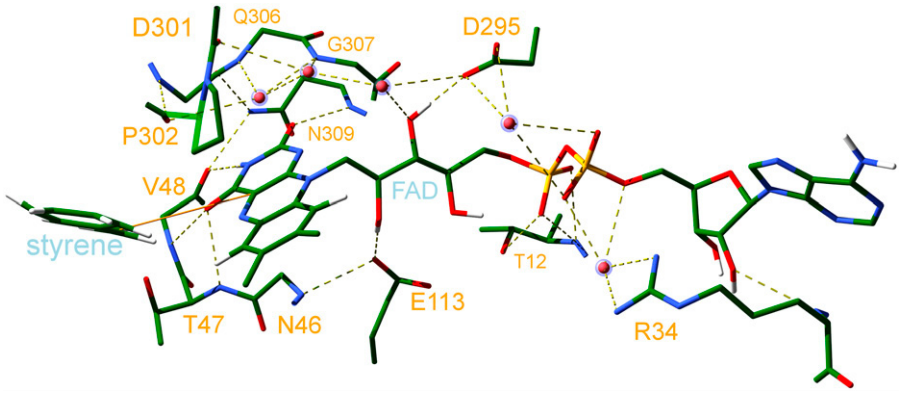


FIGURE 4 FAD and water binding in the StyA active site. Residues in the read are labeled with a smaller font. The substrate styrene is shown as well. Hydrogen bonds among FAD, selected binding residues, and selected water molecules are indicated with dashed yellow lines. Waters are highlighted with transparent light-blue spheres.

reaction (15), which suggests that the Pro^{302} and Thr^{47} are important for catalytic activity in StyA. Most discussion in literature on the importance of water molecules in the pHBH catalytic function, appears to have been focused on a hydrogen-bonded water network in the active site cavity bridging the substrate and the putative catalytic Pro^{293} with the bulk water through the FAD binding pocket (39,40). Likewise, the strongly hydrophobic StyA styrene binding cavity of the StyA model contains no water molecules (Fig. 4), but three water molecules were hydrogen-bonded to the FAD side of the P302 carbonyl and form a hydrogen-bonded network with the backbone of residues 301, 302, 306, and 307 and the flavin ribityl of FAD (Fig. 4). In the crystal structures of pHBH an arrangement of water molecules is present, involving the active-site P293 carbonyl, which is very similar to that found in the current model. It should be noted that in the StyA model the crystal waters of the pHBH template structure were excluded, and the StyA protein structure was resolvated and equilibrated in water for the MD simulations. These water molecules have been suggested to be involved in the catalytic mechanism of pHBH for proton donation as well as stabilization of intermediate state(s) (40), which leads to the compelling suggestion that also in StyA, these waters are critical for catalysis.

CONCLUSIONS

A first homology model of StyA has been constructed and optimized using molecular dynamics simulations. For all key residues for ligand and cofactor binding and for catalytic activity, known from the template pHBH protein, corresponding counterparts in the StyA sequence and structure could be identified. Known experimental data on enantioselectivity and catalytic activity of styrene in StyA wild types and mutants and of different substrates in the wild-type StyA corresponded well with predicted substrate binding orientations and affinities based on this StyA homology model. The StyA active site residues Thr^{47} and Pro^{302} are critical in StyA catalysis and are possibly involved in the stabilization of reaction intermediate(s). The remainder of the styrene binding pocket is exclusively hydrophobic and consists of Phe^{235}, Val^{303}, Leu^{220}, the di-methyl-phenyl ring of FAD, Ile^{198}, and Val^{222}, where styrene binds stacked between Phe^{235} and Val^{303}. The model structure of StyA is deposited in the PDB under ID 2HD8. The current StyA protein model thus provides a good starting point for further validation of the role of the active site residues and FAD cofactor binding residues in ligand binding and catalysis. Novel insights in the active site topology and in the binding of ligands in SMO/StyA may facilitate the rational design of mutants with specific, altered enantioselective properties.

We thank Dr. Chris Oostenbrink and Dr. Katja Otto for helpful discussions and Simon Allioth for assisting with protein purifications.

Financial support from the European Community grant ''Industrial biocatalysis with new oxygenases in a novel electro-enzyme reactor'' in the ''Quality of Life and Management of Living Resources'' program is acknowledged.

REFERENCES

1. Panke, S., M. G. Wubbolts, A. Schmid, and B. Witholt. 2000. Production of enantiopure styrene oxide by recombinant *Escherichia coli* synthesizing a two-component styrene monooxygenase. *Biotechnol. Bioeng.* 69:91–100.
2. Kantz, A., F. Chin, N. Nallamothu, T. Nguyen, and G. T. Gassner. 2005. Mechanism of flavin transfer and oxygen activation by the two-component flavoenzyme styrene monooxygenase. *Arch. Biochem. Biophys.* 442:102–116.
3. Hollmann, F., P.-C. Lin, B. Witholt, and A. Schmid. 2003. Stereospecific biocatalytic epoxidation: the first example of direct regeneration of a FAD-dependent monooxygenase for catalysis. *J. Am. Chem. Soc.* 125:8209–8217.
4. Otto, K., K. Hofstetter, M. Röthlisberger, B. Witholt, and A. Schmid. 2004. Biochemical characterization of StyAB from *Pseudomonas* sp. VLB120 as a two-component flavin-diffusible monooxygenase. *J. Bacteriol.* 186:5292–5302.
5. Hofstetter, K. 2006. Biocatalytic asymmetric epoxidation of styrene and derivatives using isolated styrene monooxygenase. PhD thesis. ETH, Zürich. http://e-collection.ethbib.ethz.ch/show?type=diss&nr=16400.
6. Schmid, A., K. Hofstetter, H.-J. Feiten, F. Hollmann, and B. Witholt. 2001. Integrated biocatalytic synthesis on gram scale: the highly enantioselective preparation of chrial oxirans with styrene monooxygenase. *Adv. Synth. Catal.* 343:732–737.
7. Entsch, B., L. J. Cole, and D. P. Ballou. 2005. Protein dynamics and electrostatics in the function of *p*-hydroxybenzoate hydroxylase. *Arch. Biochem. Biophys.* 433:297–311.
8. Entsch, B., and W. J. H. Van Berkel. 1995. Flavoprotein structure and mechanism. 1. structure and mechanism of *para*-hydroxybenzoate hydroxylase. *FASEB J.* 9:476–483.
9. Berman, H. M., J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, and P. E. Bourne. 2000. The Protein Data Bank. *Nucleic Acids Res.* 28:235–242.
10. Altschul, S. F., W. Gish, W. Miller, E. W. Meyers, and D. J. Lipman. 1990. Basic local alignment search tool. *J. Mol. Biol.* 215:403–410.
11. Schreuder, H. A., P. A. Prick, R. K. Wierenga, G. Vriend, K. S. Wilson, W. G. Hol, and J. Drenth. 1989. Crystal structure of the *p*-hydroxybenzoate hydroxylase-substrate complex refined at 1.9 A resolution. Analysis of the enzyme-substrate and enzyme-product complexes. *J. Mol. Biol.* 208:679–696.
12. Enroth, C., H. Neujahr, G. Schneider, and Y. Lindqvist. 1998. The crystal structure of phenol hydroxylase in complex with FAD and phenol provides evidence for a concerted conformational change in the enzyme and its cofactor during catalysis. *Structure.* 6:605–617.
13. Eppink, M. H., H. A. Schreuder, and W. J. Van Berkel. 1997. Identification of a novel conserved sequence motif in flavoprotein hydroxylases with a putative dual function in FAD/NAD(P)H binding. *Protein Sci.* 6:2454–2458.
14. Meller, J., and R. Elber. 2001. Linear programming optimization and a double statistical filter for protein threading protocols. *Proteins.* 45: 241–261.
15. Ridder, L., A. J. Mulholland, I. Rietjens, and J. Vervoort. 2000. A quantum mechanical/molecular mechanical study of the hydroxylation of phenol and halogenated derivatives by phenol hydroxylase. *J. Am. Chem. Soc.* 122:8728–8738.
16. Nishiya, Y., and T. Imanaka. 1996. Analysis of the interaction between the *Arthrobacter* sarcosine oxidase and the coenzyme flavin adenine dinucleotide by site-directed mutagenisis. *Appl. Environ. Microbiol.* 62:2405–2410.
17. Gatti, D., B. Pafley, M. Lah, B. Entsch, V. Maasey, D. Ballou, and M. Ludwig. 1994. The mobile flavin of 4-OH benzoate hydroxylase. *Science.* 266:110–114.

18. Sali, A., and T. L. Blundell. 1993. Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* 234:779–815.

19. Laskowski, R. A., M. W. MacArthur, D. S. Moss, and J. M. Thornton. 1993. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* 26:283–291.

20. Berendsen, H. J. C., D. van der Spoel, and R. van Drunen. 1995. Gromacs: A message-passing parallel molecular dynamics implementation. *Comput. Phys. Commun.* 91:43–56.

21. Lindahl, E., B. Hess, and D. van der Spoel. 2001. GROMACS 3.0: a package for molecular simulation and trajectory analysis. *J. Mol. Model. (Online).* 7:306–317.

22. Daura, X., A. E. Mark, and W. F. v. Gunsteren. 1998. Parametrisation of aliphatic CH_n united atoms of GROMOS96 force field. *J. Comput. Chem.* 19:535–547.

23. Van Gunsteren, W. F., S. R. Billeter, A. A. Eising, P. H. Hunenberger, P. Kruger, A. E. Mark, W. R. P. Scott, and I. G. Tironi. 1996. Biomolecular Simulation: Gromos96 Manual and User Guide. Biomos b.v., Zurich, Groningen.

24. Hess, B., H. Bekker, H. J. C. Berendsen, and J. G. E. M. Fraaije. 1997. LINCS: A linear constraint solver for molecular simulations. *J. Comput. Chem.* 18:1463–1472.

25. Amadei, A., A. B. Linssen, and H. J. Berendsen. 1993. Essential dynamics of proteins. *Proteins.* 17:412–425.

26. Hünenberger, P. H., A. E. Mark, and W. F. Van Gunsteren. 1995. Fluctuation and cross-correlation analysis of protein motions observed in nanosecond molecular dynamics simulations. *J. Mol. Biol.* 252: 492–503.

27. Summers, N. L., and M. Karplus. 1991. Modeling of side chains, loops, and insertions in proteins. *Methods Enzymol.* 202:156–204.

28. Jones, G., P. Willett, R. C. Glen, A. R. Leach, and R. Taylor. 1997. Development and validation of a genetic algorithm for flexible docking. *J. Mol. Biol.* 267:727–748.

29. Babor, M., V. Sobolev, and M. Edelman. 2002. Conserved positions for ribose recognition: Importance of water bridging interactions among ATP, ADP and FAD-protein complexes. *J. Mol. Biol.* 323:523–532.

30. Åqvist, J., C. Medina, and J. E. Samuelsson. 1994. A new method for predicting binding affinity in computer-aided drug design. *Protein Eng.* 7:385–391.

31. Hansson, T., J. Marelius, and J. Åqvist. 1998. Ligand binding affinity prediction by linear interaction energy methods. *J. Comput. Aided Mol. Des.* 12:27–35.

32. Sanner, M. F., A. J. Olson, and J. C. Spehner. 1996. Reduced surface: an efficient way to compute molecular surfaces. *Biopolymers.* 38:305–320.

33. Wang, W., J. Wang, and P. A. Kollman. 1999. What determines the van der Waals coefficient beta in the LIE (linear interaction energy) method to estimate binding free energies using molecular dynamics simulations? *Proteins.* 34:395–402.

34. Åqvist, J., and T. Hansson. 1996. On the validity of electrostatic linear response in polar solvents. *J. Phys. Chem.* 100:9512–9521.

35. Wierenga, R. K., P. Terpstra, and W. G. J. Hol. 1986. Prediction of the occurrence of the ADP-binding beta-alpha-beta-fold in proteins, using an amino-acid-sequence fingerprint. *J. Mol. Biol.* 187:101–107.

36. Kleiger, G., and D. Eisenberg. 2002. GXXXG and GXXXA motifs stabilize FAD and NAD(P)-binding Rossmann folds through C-alpha-H $\cdots$ O hydrogen bonds and van der Waals interactions. *J. Mol. Biol.* 323:69–76.

37. Hess, B. 2002. Convergence of sampling in protein simulations. *Phys. Rev. E.* 65:31911.

38. Palfey, B. A., D. P. Ballou, and V. Massey. 1997. Flavin conformational changes in the catalytic cycle of *p*-hydroxybenzoate hydroxylase substituted with 6-azido- and 6-aminoflavin adenine dinucleotide. *Biochemistry.* 36:15713–15723.

39. Ortiz-Maldonado, M., B. Entsch, and D. P. Ballou. 2004. Oxygen reactions in *p*-hydroxybenzoate hydroxylase utilize the H-bond network during catalysis. *Biochemistry.* 43:15246–15257.

40. Schreuder, H. A., W. G. Hol, and J. Drenth. 1990. Analysis of the active site of the flavoprotein *p*-hydroxybenzoate hydroxylase and some ideas with respect to its reaction mechanism. *Biochemistry.* 29:3101–3108.

41. Clamp, M., J. Cuff, S. M. Searle, and G. J. Barton. 2004. The Jalview Java Alignment Editor. *Bioinformatics.* 12:426–427.